NeuroImage 121 (2015) 91–105

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Reliability comparison of spontaneous brain activities between BOLD and CBF contrasts in eyes-open and eyes-closed resting states

Qihong Zou [a,*], Xinyuan Miao [b], Dapeng Liu [b], Danny J.J. Wang [c], Yan Zhuo [b], Jia-Hong Gao [a,d,e,*]

[a] Center for MRI Research, Peking University, Beijing, China
[b] State Key Laboratory of Brain and Cognitive Science, Beijing MRI Center for Brain Research, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China
[c] Department of Neurology, University of California, Los Angeles, CA, USA
[d] Beijing City Key Lab for Medical Physics and Engineering, School of Physics, Peking University, Beijing, China
[e] McGovern Institute for Brain Research, Peking University, Beijing, China

ARTICLE INFO

Article history:
Received 8 January 2015
Accepted 17 July 2015
Available online 27 July 2015

Keywords:
Oscillation
Cerebral blood flow
Resting state
Eyes open
Eyes closed

ABSTRACT

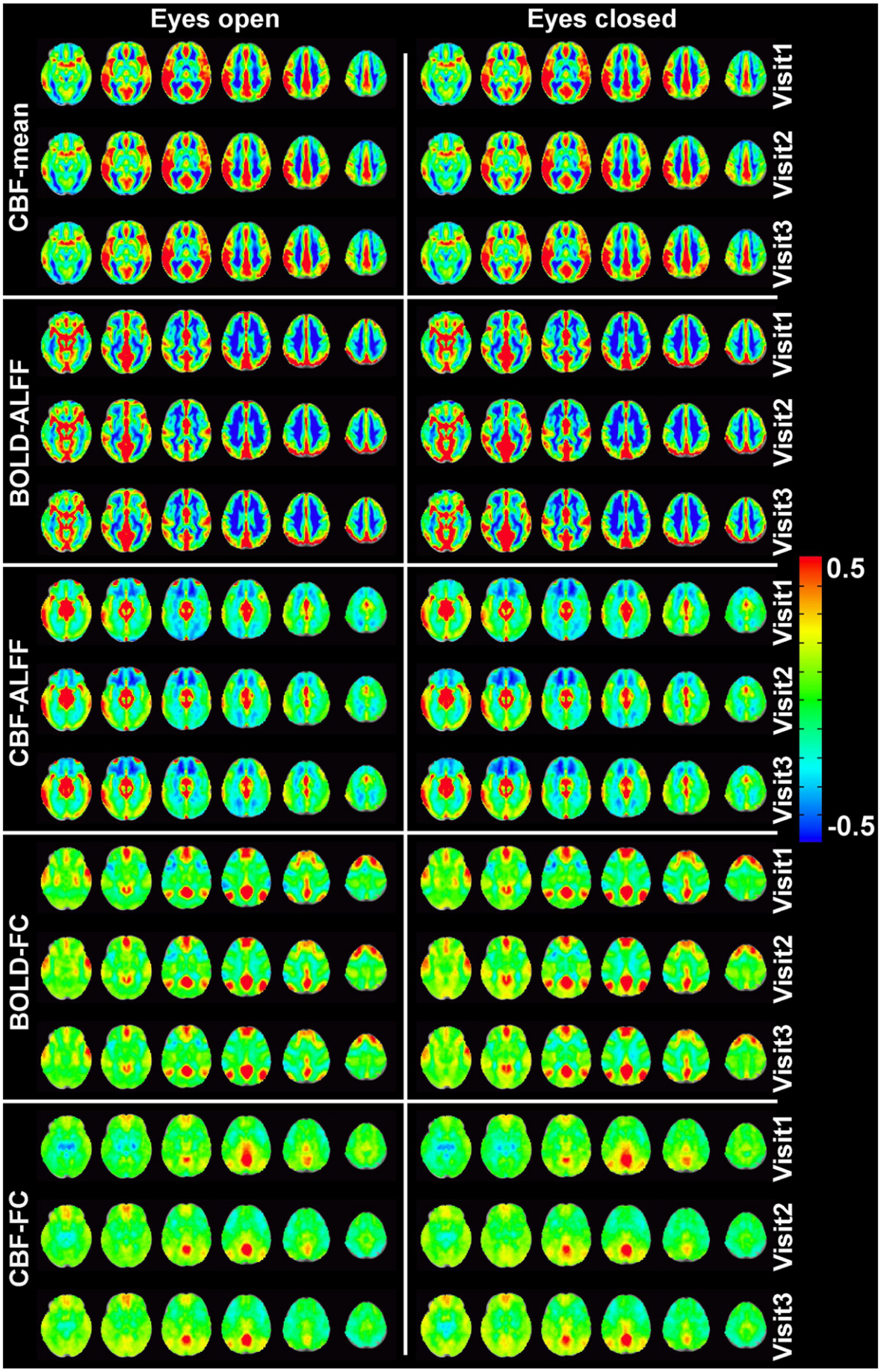
Blood oxygenation level dependent (BOLD) and arterial spin labeling (ASL) are two predominant resting-state fMRI techniques in mapping spontaneous brain activities. At single voxel level, cerebral blood flow (CBF) measured by ASL and amplitude of low frequency fluctuations (ALFF) of BOLD have been recognized as useful indices to characterize brain function in health and disease. However, no study has directly compared the test–retest reliability between BOLD and CBF contrasts on the same group of subjects at single voxel level. Moreover, both eyes-open and eyes-closed conditions have been employed as resting states, but it is still not clear which state is more reliable. Here we collected BOLD and ASL data during eyes-open and eyes-closed states across three scanning visits on twenty-two healthy young subjects. CBF-mean, BOLD- and CBF-ALFF were computed to characterize corresponding brain activities at single voxel level. Seed-based functional connectivity (FC) with the posterior cingulate cortex (PCC) was further calculated for both BOLD and CBF data. Intra-class correlation was used as the index of long-term reliability between visits 1 and 2 (two months apart) and short-term reliability between visits 2 and 3 (on the same day). Both short- and long-term reliabilities for CBF-mean and BOLD-ALFF were high, but were lower for CBF-ALFF, BOLD- and CBF-FC in both eyes-open and eyes-closed states. Direct comparisons showed that brain regions with the highest reliability of CBF-mean were mainly in the gray matter. The reliability of CBF-ALFF and BOLD-FC was lower than that of BOLD-ALFF, and the reliability of CBF-FC was lower than those of both CBF-ALFF and BOLD-FC. Furthermore, we observed that reliabilities of the eyes-open state were higher than those of the eyes-closed state for both imaging contrasts, though the effect size was small. Voxel-wise comparisons demonstrated that the long-term reliability of BOLD-ALFF was significantly higher with eyes open in the visual system, and both the short- and long-term reliability of BOLD-FC was slightly higher with eyes open in the default mode network. Moreover, we showed that denoising decreased the reliability of both ALFF and FC of both BOLD and ASL contrasts. In conclusion, our results indicated that CBF-mean and BOLD-ALFF could both be used as reliable indices for characterizing resting-state brain activities at single voxel level and recommended the eyes-open state for resting-state studies, especially for those targeting the visual system and default mode network.

Introduction

Credibility and reproducibility are everything for science (Marcus, 2014). Thus they are of crucial importance to evaluate the test–retest reliability of brain activity measures (Hodkinson et al., 2013; Jahng et al., 2005; Jann et al., 2014; Kublbock et al., 2014; Patriat et al., 2013; Shehzad et al., 2009; Zhu et al., 2014; Zuo et al., 2010a,b,c, 2013; Zuo and Xing, 2014). Resting-state fMRI, including blood oxygenation level dependent (BOLD) and arterial spin labeling (ASL), has become a powerful tool in mapping human brain functions (Biswal et al., 2010; Smith et al., 2013). Previous investigations on test–retest reliability of spontaneous BOLD signal oscillations (Kublbock et al., 2014; Zuo et al., 2010a) showed that regions with high reliability were mainly located in the gray matter. Similar results were observed by separate studies investigating test–retest reliability of average CBF (Hodkinson et al., 2013; Jahng et al., 2005; Jann et al., 2015). One recent study compared short-term test–retest reliability of CBF-mean, functional connectivity based on BOLD and CBF time series, which showed generally more reliable spatial patterns of functional connectivity of BOLD than that of CBF, while the highest reproducibility was yielded by CBF-mean (Jann et al., 2015). However, functional connectivity provides measures of inter-regional relationship, i.e., relational characteristics between brain regions, while mean CBF reflects static characteristics of a region, i.e., regional characteristics at single voxel level (Zuo and Xing, 2014).

* Corresponding authors at: Center for MRI Research, Peking University, Beijing 100871, China.
E-mail addresses: zouqihong@pku.edu.cn (Q. Zou), jgao@pku.edu.cn (J.-H. Gao).

http://dx.doi.org/10.1016/j.neuroimage.2015.07.044
1053-8119/


Eyes open
Eyes closed
CBF-mean
BOLD-ALFF
CBF-ALFF
BOLD-FC
CBF-FC
Visit1
Visit2
Visit3
0.5
-0.5

It is not straightforward to compare reliability of a relational metric among many voxels to a regional one at the single voxel level. In contrast to functional connectivity, amplitude of low-frequency fluctuation (ALFF) is a metric that measures the fluctuation amplitude of each single voxel (Liu et al., 2013; Zang et al., 2007; Zou et al., 2008; Zuo and Xing, 2014) and could be applied to both BOLD and ASL techniques. ALFF has been widely used to examine moment-to-moment brain signal variability (Garrett et al., 2011; Garrett et al., 2013), to correlate with task activations and behavior (Zou et al., 2013), and to characterize disease stages (Han et al., 2011; Lui et al., 2009; Wang et al., 2011b; Yu et al., 2014). Direct comparison of the short- and long-term test–retest reliability of BOLD and CBF contrasts on an identical group of subjects at single voxel level is still lacking. Thus the first goal of the current study is to fill this gap.

Furthermore, there are no consistent guidelines on how to perform a resting-state scan across various studies (Tagliazucchi and Laufs, 2014). Subjects are usually asked to not think of anything in particular and keep eyes open (EO) or eyes closed (EC). However, it has been widely shown that EO and EC are two physiological states with different levels of brain activity. It was observed almost one century ago by Berger (Berger, 1929, 1930) that the EEG alpha rhythm disappeared when the eyes were open, as compared to the case when the eyes were closed. BOLD fMRI studies have shown brain activity differences between the EO and EC states in the visual cortex (Jao et al., 2013; Liu et al., 2013; Yang et al., 2007; Zou et al., 2009; Zou et al., 2015), motor and auditory cortices (Jao et al., 2013; Liu et al., 2013; Yuan et al., 2014; Zou et al., 2015), and the default mode network (Jao et al., 2013; Liu et al., 2013; Yan et al., 2009). Using the ASL technique, significant CBF increases in the primary and secondary visual areas (Brodmann area (BA) 17, 18) for EO compared with EC state (Hermes et al., 2007; Zou et al., 2015) have been reported. Thus inconsistencies in resting states across studies may confound comparisons and interpretations among varied studies and impede big-data sharing which is of growing interest. Therefore, our second goal is to compare the reliability of resting-state activity during EO and EC states.

Moreover, it has been demonstrated that denoising strategies such as noise components removal using independent component analysis (ICA) and regression of physiological noise could boost the predictability of age using BOLD-ALFF (Garrett et al., 2010) and leading to more accurate estimates of group maps of BOLD-FC (Birn et al., 2014). However, physiological pulsations (Birn et al., 2014) and motion (Van Dijk et al., 2012; Zuo et al., 2014) showed moderate reliability, which might inflate the reliability of ALFF and FC of BOLD data. Our third goal is to compare the reliability of resting-state activity with and without data denoising during preprocessing.

To this end we collected resting-state fMRI data using BOLD and ASL in both EO and EC states across three scanning visits, with two visits performed on the same day and the rest one performed two months beforehand. Intra-Class Correlation (ICC) (Shrout and Fleiss, 1979), a well-established metric of test–retest reliability, was used. The two visits on the same day were used for assessing short-term reliability, while the first and second visits that were two month apart were used for assessing long-term reliability. Reliabilities between two imaging contrasts were compared qualitatively and quantitatively. Similar comparisons on reliabilities between two resting states and different preprocessing strategies were conducted.

Materials and methods

Subjects

Twenty-two healthy volunteers (24.4 ± 2.2 years old, 12 females) participated in the study. They were recruited following a protocol approved by the IRB of our institution. All subjects were screened with a questionnaire to ensure they had no history of neurological diseases or psychiatric disorders. Signed informed consent was obtained from all subjects prior to study enrollment.

Experimental paradigm

Each subject was imaged for three scanning visits. Visits 2 and 3 were conducted on the morning and afternoon of the same day, which were about two months (74.7 ± 7.7 days) following visit 1. For each scanning visit, four resting-state scans, i.e., both EC and EO with a black screen using both BOLD and ASL techniques were acquired for each subject. The order of the two resting states and the two fMRI contrasts was counter-balanced across subjects. In total, 12 resting-state scans were acquired for each subject across three scanning visits, with two resting states (EC and EO) and two imaging contrasts (BOLD and ASL).

Data acquisition

MRI data were collected on a 3 T Siemens MR scanner (Siemens, Erlangen, Germany) using the body coil for transmission and a 12-channel head coil for reception. Foam pads were used to restrain head motions and scanner noise was attenuated by earplugs.

A pseudo continuous ASL (pCASL) technique (Dai et al., 2008; Wu et al., 2007) was used for CBF measurement. Interleaved control and label images were acquired using a single-shot 3D GRASE sequence with the following parameters: TR/TE = 3000/23.38 ms, slice thickness = 5 mm without gap, 24 slices, FOV = 220 × 220 mm^2 with in-plane resolution of 3.44 × 3.44 mm^2, GRAPPA factor = 2, partial Fourier = 6/8, labeling duration = 1104ms, label offset = 90 mm, post-labeling delay = 1000 ms and global inversion pulses for background suppression (Jann et al., 2014). The duration of the resting-state pCASL scan was 8 min, and it included 80 pairs of control and label images.

A single-shot gradient echo EPI sequence was used to acquire BOLD data. For comparison, most acquisition parameters of the BOLD were kept the same as those of pCASL, except for TR/TE (2500/30 ms), FA (80°), GRAPPA factor (no parallel imaging was used), and partial Fourier (not used). The duration of the resting-state BOLD scan was 8 min, and it produced 192 images.

For registration purposes, high-resolution anatomical images were acquired using a 3D magnetization-prepared rapid gradient echo T1-weighted sequence (TI/TR/TE = 1100/2530/3.37 ms, flip angle = 7°, 176 slices, FOV = 256 × 256 mm^2, 1 × 1 × 1 mm^3 resolution) on each subject.

Data processing

Both resting-state pCASL and BOLD data were preprocessed using the Analysis of Functional Neuroimages (AFNI) software package (Cox, 1996), FSL (http://www.fmrib.ox.ac.uk/fsl) and MATLAB 7.1 (The Mathworks, Inc.). Identical processing pipelines were adopted for data acquired during EO and EC states.

CBF-mean

For pCASL data, control and label images with the least deviation from the trend of the respective time courses were chosen as the base for motion correction and were linearly registered to each other (Liang et al., 2013; Zou et al., 2011). Then, control and label images were registered to their corresponding bases, separately. It is noted that slice time correction was not performed for pCASL data with 3D

Fig. 1. Group average maps of five metrics at three visits in two resting states. For indices CBF-mean, BOLD- and CBF-ALFF, a hot color represents higher values than the global mean, and a cold color represents lower than the global mean. For indices BOLD- and CBF-FC, a hot color represents positive correlation with the posterior cingulate cortex (PCC), and a cold color represents negative correlation with the PCC.

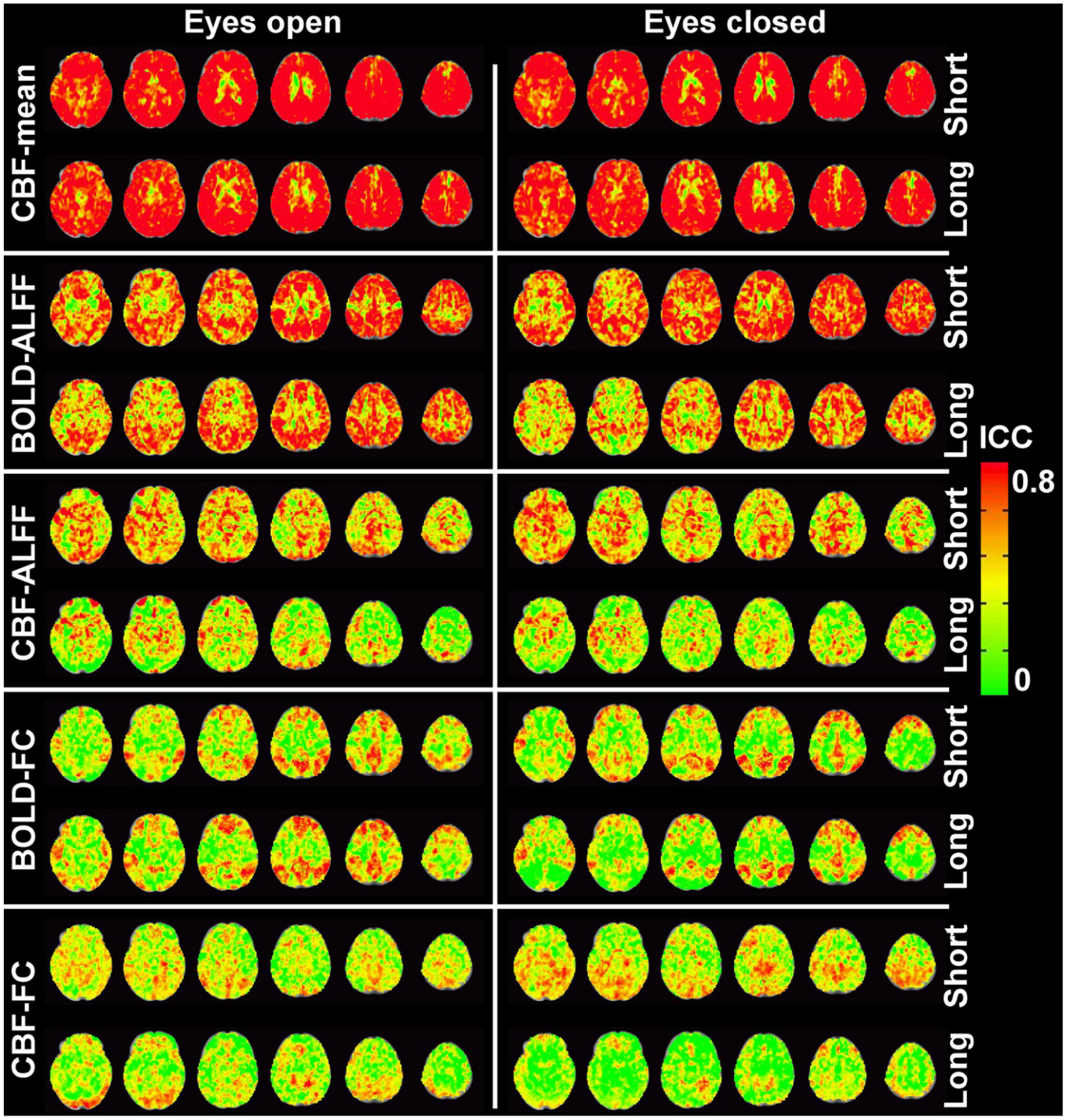


Fig. 2. Short- and long-term ICC maps. ICC maps of CBF-mean, BOLD- and CBF-ALFF, and BOLD- and CBF-FC indices computed from two imaging contrasts in EO (left columns) and EC (right columns) states. "Short" indicates short-term ICC while "long" indicates long-term ICC. *Z*-axial coordinates in the Talairach and Tournoux space are from −9 to 51 mm in steps of 12 mm.

acquisition. Following head motion correction, label images were sinc subtracted from control images to obtain CBF-weighted time series while minimizing BOLD contaminations (Chuang et al., 2008). The CBF-weighted images were spatially smoothed with a 6-mm Gaussian kernel along the x and y directions only, as there was spatial smoothing along *z* direction with 3D acquisition. Then, the spatially smoothed CBF-weighted time series were spatially normalized to Talairach and Tournoux space with a resampling resolution of $3 \times 3 \times 3$ mm^3 and averaged across the time series to obtain a CBF-mean map. A relative CBF map was further obtained by dividing the CBF-mean map with the mean CBF of the entire brain voxels.

BOLD-ALFF and FC

For BOLD data, preprocessing steps included slice-timing correction, head motion correction, linear trend removal, spatial smoothing with a 6-mm Gaussian kernel along all three directions, and spatial

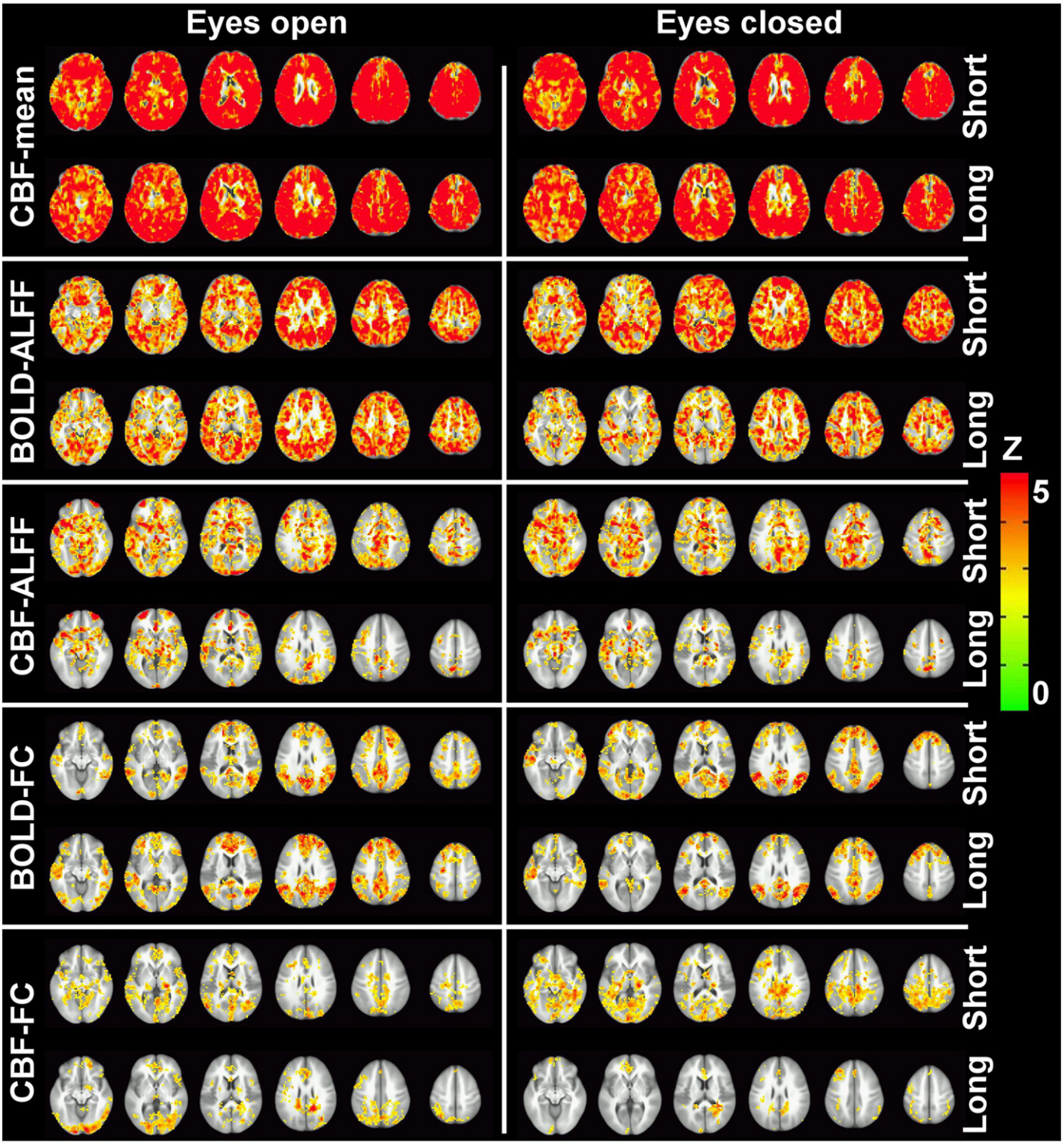


Fig. 3. Significance of the short- and long-term ICC maps. Significant ICC maps of five indices computed from two imaging contrasts in two resting states (display and organization are analog to Fig. 2). *Z*-axial coordinates in the Talairach and Tournoux space are from −9 to 51 mm in steps of 12 mm.

normalization to Talairach and Tournoux space with a resampling resolution of $3 \times 3 \times 3$ mm^3. Physiological noises and head motion are well-known non-neuronal artifacts to fMRI data. To eliminate the potential contamination of physiological noise including cardiac and respiratory pulsations in the BOLD data, we adopted a validated temporal independent component analysis (ICA) methodology at single subject level. Physiologic EStimation by Temporal ICA or PESTICA, detects the pulse and the breathing cycles from the individual BOLD data itself, which has been demonstrate to be equivalent to a parallel monitored pulse signal and a respiratory chest-bellow signal (Beall and Lowe, 2007, 2010). To reduce the potential effect of head motion, we performed further analysis to regress out the six motion parameters. In addition, we regressed out the average white matter and cerebrospinal fluid (CSF) signal which were generally considered as covariates of no interest. The residual time series were used for ALFF and FC analyses. We calculated voxel-wise BOLD-ALFF over a frequency range of 0.01–0.083 Hz. The BOLD-ALFF maps were

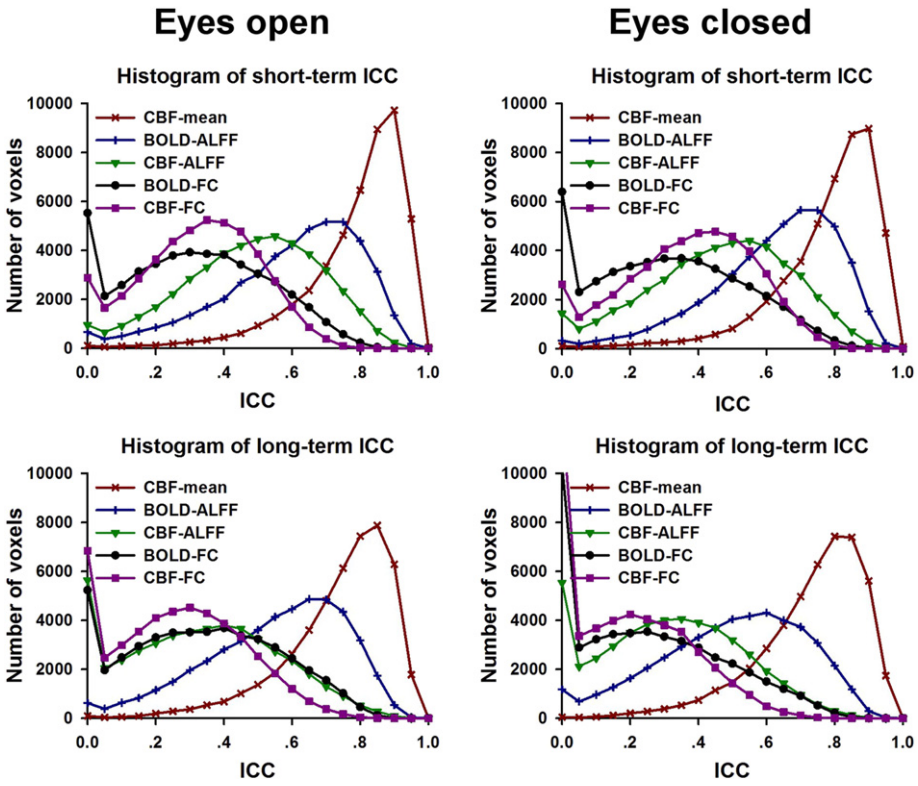


Fig. 4. Histograms of short- and long-term ICC among different metrics. Histograms of short- (top row) and long-term (bottom row) ICC among the five metrics including CBF-mean, BOLD- and CBF-ALFF, and BOLD- and CBF-FC in EO (left column) and EC (right column) states.

standardized by normalization with the mean BOLD-ALFF of the entire brain voxels (Zou et al., 2008; Zuo et al., 2010a). Prior to FC analysis, the residual BOLD time series were band-pass filtered (0.01–0.083 Hz) and a 10-mm diameter spherical seed located in the posterior cingulate cortex (PCC) was selected (Van Dijk et al., 2012; Yan et al., 2013b). We calculated the Pearson's correlation coefficient between the seed time series and every other voxel in the brain and transformed the correlation coefficient values into *z* values using Fisher's *r*-to-*z* transformation. To evaluate the effect of data denoising, we calculated BOLD-ALFF with standard preprocessing, i.e., with neither ICA denoising nor regression of covariates of no interest (Zou et al., 2008; Zuo et al., 2010a); and BOLD-FC with standard preprocessing, i.e., without ICA denoising but with regression of covariates of no interest (Patriat et al., 2013).

CBF-ALFF and FC

The ASL data were collected using a 3D sequence, which were acquired over a much longer timeframe than that of 2D BOLD data. The effect of physiological noise is much harder to correct for 3D data, as the physiological noise might be largely smeared during longer 3D acquisition. So far as we know there is no validated methodology for physiological noise correction for fMRI data collected using 3D sequences, thus physiological noise modeling was not applied to ASL data. Nevertheless, regression of covariates of no interest including head motion and average signal of white matter and CSF was performed on the ASL time series that were preprocessed as described above (CBF-mean section). Similar to BOLD-ALFF, we calculated voxel-wise CBF-ALFF over a frequency range of 0.01 to 0.083 Hz, which were standardized by normalization with the mean CBF-ALFF of the entire brain voxels (Zou et al., 2008; Zuo et al., 2010a). Similar to BOLD-FC, CBF-FC with the same PCC seed was calculated based on the band-pass filtered (0.01–0.083 Hz) residual CBF time series.

Table 1
Effect size of comparisons of ICC between different metrics.

	Eyes open		Eyes closed	
	Short-term	Long-term	Short-term	Long-term
CBF-mean vs. BOLD-ALFF	0.48	0.48	0.45	0.51
CBF-mean vs. CBF-ALFF	0.55	0.58	0.55	0.59
CBF-mean vs. BOLD-FC	0.61	0.60	0.61	0.60
CBF-mean vs. CBF-FC	0.60	0.61	0.59	0.61
BOLD-ALFF vs. CBF-ALFF	0.27	0.42	0.36	0.38
BOLD-ALFF vs. BOLD-FC	0.54	0.46	0.55	0.47
CBF-ALFF vs. CBF-FC	0.37	0.16	0.23	0.29
BOLD-FC vs. CBF-FC	0.03	0.16	0.15	0.15

Effect size was calculated from Wilcoxon's signed rank test (see Eq. (5)).

Statistical analyses

Test–retest reliability quantified by intra-class correlation

To examine the short- (within a day, between visits 2 and 3) and long-term (two months apart, between visits 1 and 2) test–retest reliability of the single voxel level CBF-mean, CBF- and BOLD-ALFF, we conducted statistical reliability analysis using the ICC, which is defined as the proportion of variability across subjects relative to the total variability in the data (Shrout and Fleiss, 1979).

$$ICC = \frac{MS_b - MS_w}{MS_b + (k-1)MS_w} \quad (1)$$

where MS_b is between-subject mean square, MS_w is within-subject mean square, k is the number of observations made on each subject, and $k = 2$ for the ICC calculation. The higher the ICC value, the lower the within-subject variance relative to between-subject variance, which indicates higher reliability. To avoid negative values and get more accurate estimation, ICC were calculated based on linear mixed-

effects modeling with restricted maximum likelihood estimates (Zuo et al., 2013), which prescribes maximization over the parameter space and variance components are non-negative (for details, see the book by (Searle et al., 1992)).

Significance of test–retest reliability using one contrast in resting state

The significance of ICC value for each voxel over the whole brain was estimated by transforming the ICC value to *Z* score (McGraw and Wong, 1996). According to Mcgraw and Wong's formula, we first transformed ICC value into *z*:

$$z = \frac{1}{2}\ln\frac{1+(k-1)ICC}{1-ICC} \tag{2}$$

where k is the number of observations made on each subject, and $k = 2$ for the ICC significance calculation. The mean of the above statistic is zero and the variance (McGraw and Wong, 1996) is:

$$\sigma^2 = \frac{k}{2(N-2)(k-1)} \tag{3}$$

where N is the number of subjects. Then we transformed z into Z scores with standard normal distribution:

$$Z = \frac{z-0}{\sqrt{\frac{k}{2(N-2)(k-1)}}} = \frac{\sqrt{N-2}}{2}\ln\frac{1+ICC}{1-ICC} = \sqrt{N-2}\tanh^{-1}(ICC). \tag{4}$$

Comparison of test–retest reliability between imaging contrasts

Histograms of ICC values across the whole brain were used to qualitatively compare the two imaging contrasts. We then conducted a series of paired Wilcoxon signed-rank tests (Yan et al., 2013b) to investigate whether the distributions of ICC values are significantly different between the resting-state activity metrics across the whole brain. The effect size was calculated as:

$$r = \frac{Z}{\sqrt{N}} \tag{5}$$

where Z is the Z-statistics from Wilcoxon's signed rank test and N is the number of observations, here $N = 2 * \text{number of voxels}$ (Yan et al., 2013b). Further in a voxel-wise fashion, similar to the significance test of ICC, we compared ICC values between two imaging contrasts using Z test as the following (Zhu et al., 2014; Zuo et al., 2013).

$$Z = \sqrt{N-2} \times \tanh^{-1}(ICC1 - ICC2) \tag{6}$$

where N is the number of participants, and *ICC1* and *ICC2* are the ICC values of the two imaging contrasts.

Comparison of test–retest reliability between resting states

To compare reliability of resting-state brain activity between EO and EC states, reliability histograms of EO and EC were plotted simultaneously. The Wilcoxon signed-rank tests were performed between EO and EC across the brain voxels. Using voxel-wise Z test, reliabilities in the two resting states are further compared.

Comparison of test–retest reliability between different preprocessing strategies

To compare reliability of resting-state brain activity between different preprocessing strategies for BOLD contrast, reliability histograms of advanced (ICA denoising and regression of covariates of no interest) and standard preprocessing (None for ALFF and regression of covariates of no interest for FC) were plotted simultaneously. The Wilcoxon signed-rank tests were performed between different preprocessing strategies across the brain. Using voxel-wise Z test, reliabilities between the two preprocessing strategies are further compared.

Multiple comparison corrections

For whole brain multiple comparison correction, significance thresholds for Z tests of ICC for each resting state, ICC differences between each pair of resting-state activity indices, and ICC differences between two states, and two preprocessing strategies, were set at corrected $p < 0.05$ based on Monte Carlo simulations. The corrected threshold corresponded to a single voxel's $p < 0.01$ and a minimum cluster size of 2862 mm^3.

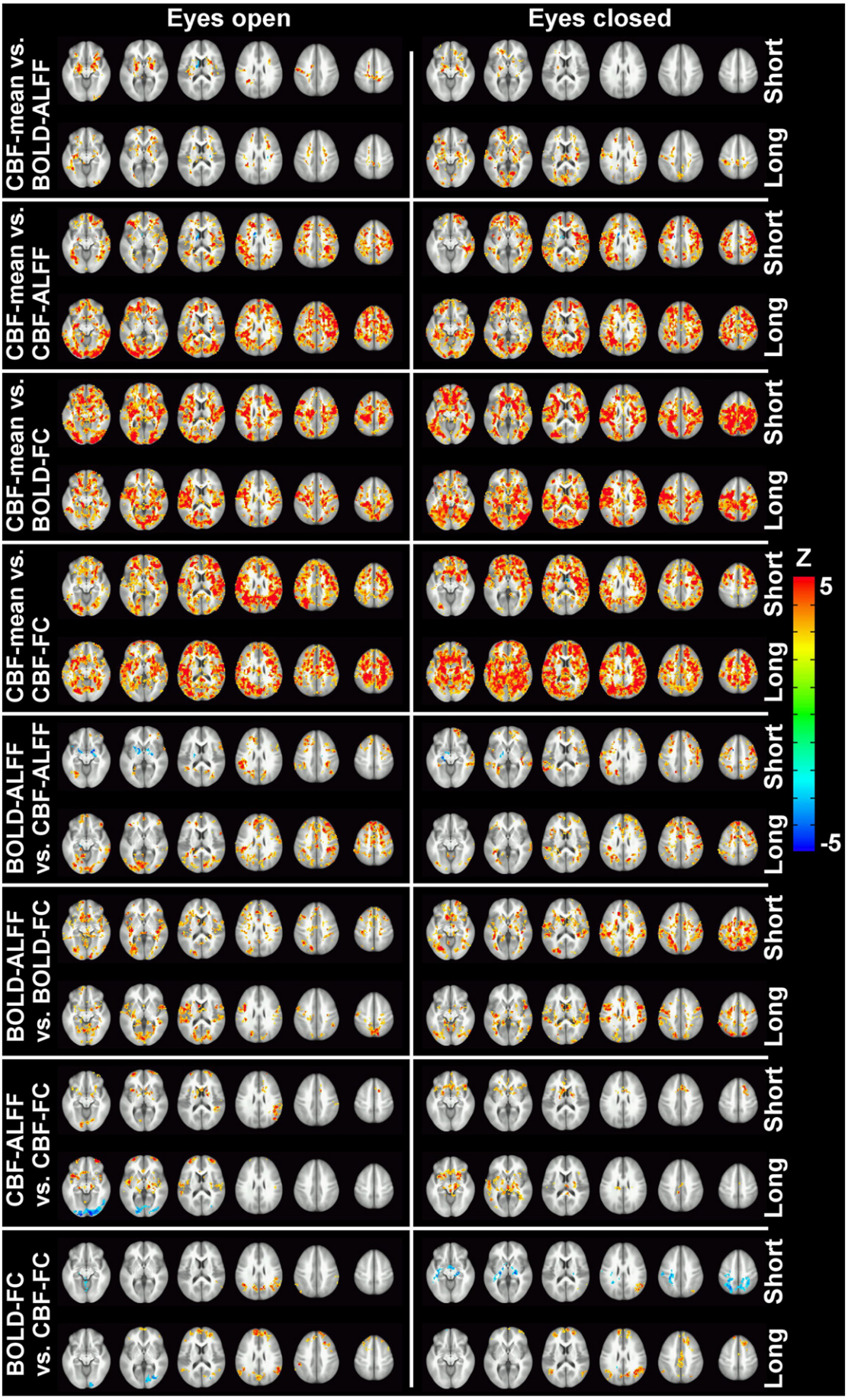
Results

Group average maps of CBF-mean, BOLD- and CBF-ALFF, and BOLD- and CBF-FC

For each resting-state activity metric, group average maps were highly similar across the three scanning visits and between the two resting states (Fig. 1). CBF-mean, BOLD- and CBF-ALFF were higher than global mean in most gray matter regions while lower than global mean in the white matter, which yielded significant spatial correlations (Table S1) among CBF-mean, BOLD- and CBF-ALFF across all the three scanning visits and between the two resting states. Based on both BOLD and ASL time series, the PCC was significantly correlated with brain regions within the default mode network such as the medial prefrontal cortex and bilateral inferior parietal lobule, while negatively correlated with brain regions in the task positive network such as lateral fronto-parietal cortices. As expected, consistently high spatial correlations were observed between BOLD- and CBF-FC. To note, BOLD- and CBF-ALFF were sensitive to physiological noises and were higher in the regions near large vessels and CSF. Although advanced preprocessing strategy including ICA noise removal (Beall and Lowe, 2007, 2010) and additional regression of head motion, CSF and white matter time series (Garrett et al., 2010), could help to eliminate noise contaminations, the observations of residual noises suggests more work on denoising is needed such as physiological noise modeling using simultaneously recorded physiological pulsations and more robust indices for measurement of fluctuation amplitude.

Test–retest reliability of CBF-mean, BOLD- and CBF-ALFF, and BOLD- and CBF-FC

For CBF-mean, consistent with previous voxel-wise test–retest studies (Hodkinson et al., 2013; Jahng et al., 2005; Jann et al., 2014), we showed that both short- and long-term ICC were excellent (ICC > 0.6) for almost the entire brain volume, especially in the gray matter in both EO and EC states (Figs. 2, 3 and S1). Consistent with previous findings (Kublbock et al., 2014; Zuo et al., 2010a) which instructed participants to keep eyes open, we observed that both short- and long-term ICC was reasonably high across the gray matter (ICC > 0.4) for BOLD-ALFF (Figs. 2, 3 and S1). In addition, we showed that ICC values were also high when the subjects kept eyes closed, except that the long-term ICC of the EC state was lower in the visual cortex. In both EO and EC states, the brain areas with high ICC values were mainly located in the gray matter, while those with relatively lower ICC values were distributed in the white matter. For CBF-ALFF, short-term ICC was observed to be reasonably high (ICC > 0.4) mainly in the cortical regions and near ventricles and large vessels, while long-term ICC was slightly lower (Figs. 2, 3 and S1). For BOLD-FC, both short- and long-term ICC were reasonably high (ICC > 0.4) for both states within the default mode network that showed positive correlations with the PCC and the task positive network that showed negative correlations with the seed. For CBF-FC, high short-term ICC values were observed to be mainly within the default mode network. The long-term ICC values were lower (Figs. 2, 3 and S1).


Eyes open
Eyes closed
CBF-mean vs. BOLD-ALFF
CBF-mean vs. CBF-ALFF
CBF-mean vs. BOLD-FC
CBF-mean vs. CBF-FC
BOLD-ALFF vs. CBF-ALFF
BOLD-ALFF vs. BOLD-FC
CBF-ALFF vs. CBF-FC
BOLD-FC vs. CBF-FC
Short
Long
Z
5
-5

Test–retest reliability between imaging contrasts

We compared reliability of the five resting-state activity indices computed using the two imaging contrasts, including 1) between CBF-mean and each of the rest four indices; 2) between ALFF and FC of the same contrasts, i.e., between BOLD-ALFF and FC, and between CBF-ALFF and FC; and 3) between imaging contrasts for ALFF and FC, i.e., between BOLD- and CBF-ALFF, and between BOLD- and CBF-FC. Histograms of ICC of CBF-mean, BOLD- and CBF-ALFF, and BOLD- and CBF-FC showed that ICC values were the highest for CBF-mean, high for BOLD-ALFF, and the lowest for CBF-ALFF, BOLD- and CBF-FC (Fig. 4) for both short- and long-term observations and in both resting states.

The Wilcoxon signed-rank tests showed significant results for all comparisons for both states and both short- and long-term ICC values (all $ps < 0.0001$). Higher ICC of CBF-mean than the rest four indices were observed with large effect size (Table 1). Higher ICC of BOLD-ALFF than CBF-ALFF and BOLD-FC were also observed with medium to large effect size (Table 1). CBF-FC showed lower ICC than CBF-ALFF and BOLD-FC, though the effect sizes were small.

Furthermore, voxel-wise *Z* tests showed that both short- and long-term ICC of CBF-mean was higher than that of BOLD-ALFF in the prefrontal cortex, parahippocampus, putamen, and supplementary motor area in both EO and EC, and higher long-term ICC additionally in the visual cortices in EC (Fig. 5). It is noted that, although the histogram of ICC showed marked difference between CBF-mean and BOLD-ALFF, voxel-wise comparisons between these two indices only led to sparsely located brain regions even when the results were uncorrected except for long-term ICC in the EC state (Fig. S2). The reliability of CBF-mean was higher than those of CBF-ALFF, BOLD- and CBF-FC in widespread cortical and subcortical regions (Figs. 5 and S2), for both the short- and long-term ICC. ICC values of BOLD- and CBF-ALFF were slightly higher than those of CBF-mean and located sparsely near the lateral ventricles.

The reliability of BOLD-ALFF was consistently higher than those of CBF-ALFF and BOLD-FC in most cortical areas including frontal, parietal, occipital and temporal regions in both EO and EC states, except that short-term reliability of CBF-ALFF was higher than BOLD-ALFF sparsely in the basal ganglia. Both the short- and long-term reliability of CBF-ALFF was significantly higher than that of CBF-FC in the basal ganglia, anterior insula, supplementary motor area and prefrontal regions. BOLD-FC showed higher reliability than CBF-FC in the key regions of the default mode network including the PCC, inferior parietal lobule and frontal regions. It should be noted that the long-term reliability of CBF-FC was higher than that of CBF-ALFF and BOLD-FC in the visual cortex in the EO state, and short-term reliability of CBF-FC was higher than BOLD-FC in the basal ganglia and parietal cortices in the EC state (Figs. 5 and S2).

Test–retest reliability of resting-state activity between EO and EC

To compare the test–retest reliability of resting-state brain activity in different states, we first compared the histograms of ICC in EO and EC states (Fig. 6). In general, ICC in EO was higher than that in EC. The histogram of ICC in the EO state only slightly shifted rightward relative to EC for CBF-mean, BOLD- and CBF-ALFF. One exception was the long-term ICC of BOLD-ALFF, the histogram of which shifted rightward considerably (Fig. 6). For both BOLD- and CBF-FC, histogram of long-term ICC in EO shifted rightward, similar to that of BOLD-ALFF. Histogram of short-term ICC for BOLD-FC was almost identical for EO and EC, while histogram of short-term ICC for CBF-FC in the EO state shifted leftward relative to EC. Wilcoxon signed-rank tests showed higher short- and long-term ICC of EO than that of EC using CBF-mean, BOLD- and CBF-ALFF, and BOLD- and CBF-FC (all $ps < 0.0001$), except for short-term ICC of BOLD-FC ($p > 0.0001$) and short-term ICC of CBF-FC (lower in EO than in EC). To note, the effect sizes between EO and EC were close to the medium for long-term ICC of BOLD-ALFF and BOLD-FC (Table 2).

We then directly evaluated the effect of resting states on ICC using voxel-by-voxel *Z* tests across the whole brain. For BOLD-ALFF, we found that the long-term ICC in EO was much higher in the primary visual cortex than that in EC (Fig. 7). The long-term ICC of CBF-ALFF in the lateral prefrontal cortex was also higher in EO than in EC (Fig. 7). The short-term ICC of BOLD-FC was higher than EC in the right superior parietal gyrus, whereas the long-term ICC of BOLD-FC was higher in the medial prefrontal cortex and bilateral higher-order visual cortices in EO. Resting states showed minimal effect on ICC of CBF-mean and CBF-FC, and short-term ICC values of BOLD- and CBF-ALFF, even when the results were uncorrected for multiple comparisons (Fig. S3). To note, though histogram of CBF-FC showed higher short-term ICC in EC, state effect was observed to be negligible using voxel-wise analysis. In general, there was a trend of increased ICC of BOLD- and CBF-ALFF, and BOLD- and CBF-FC in the visual, motor and auditory cortices and key default mode network regions such as precuneus in the EO relative to the EC state (Fig. S3). Interestingly, the short-term ICC of BOLD-ALFF tended to decrease in the primary motor and auditory cortices and thalamus, whereas the short-term ICC of BOLD-FC tended to decrease in the bilateral higher-order visual cortices with eyes open.

Test–retest reliability of resting-state activity between different preprocessing strategies

Histograms of ICC BOLD-ALFF, and BOLD-FC with data denoising (advanced preprocessing) consistently shifted leftward compared to those of ICC without data denoising (standard preprocessing) (Fig. 8), indicating reduced reliability of resting-state activity with data denoising. Wilcoxon signed-rank tests showed significant effect of preprocessing strategies on both short- and long-term ICC for all indices in two resting states with small to medium effect sizes (Table 3). Using voxel-wise *Z* tests, in general, the effect of data denoising on ICC values of BOLD-ALFF and BOLD-FC was minimal though with a decreasing trend (Fig. S4).

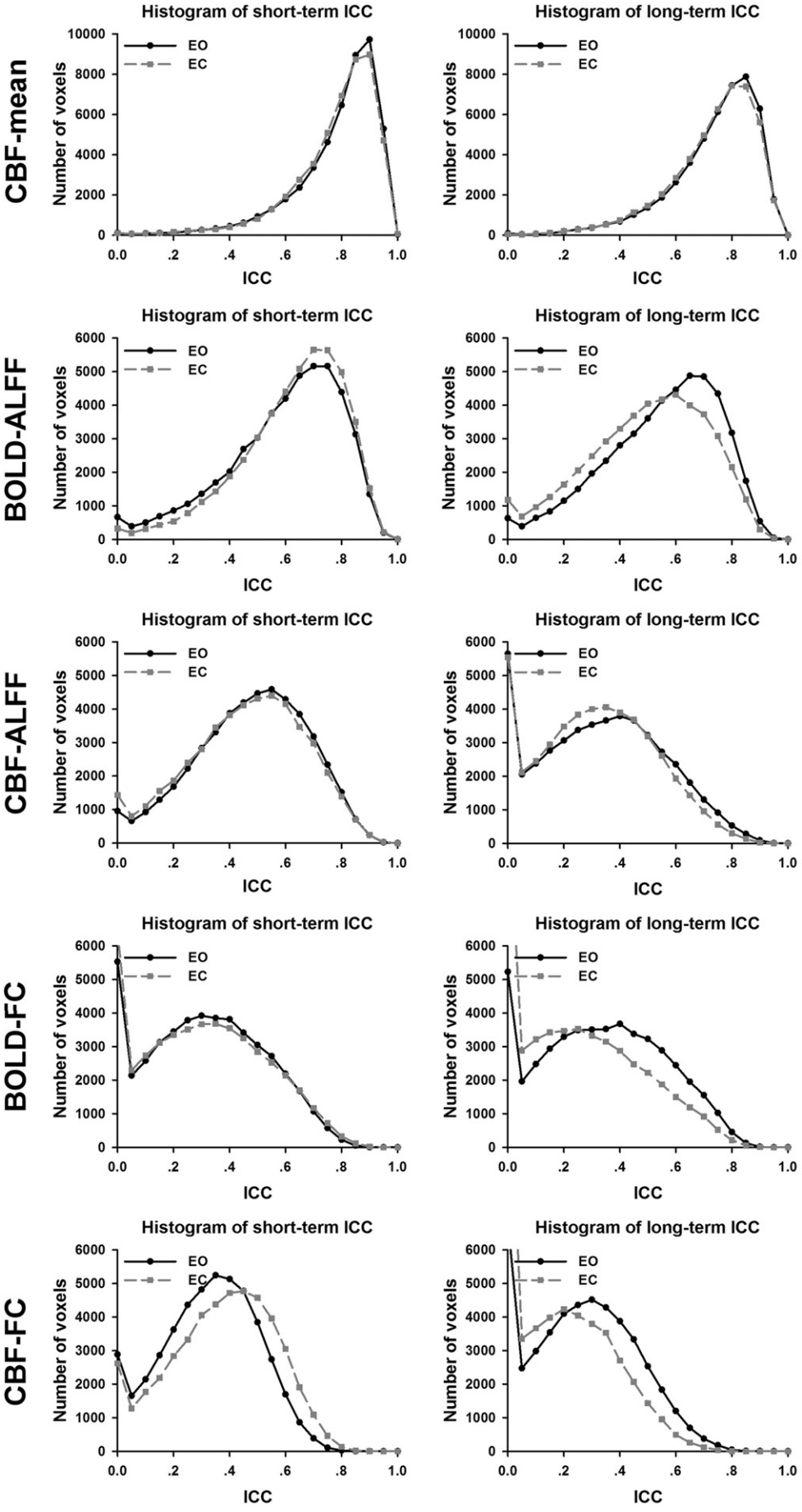
Discussion

In the current study, we compared the short- (within a day) and long-term (two months apart) test–retest reliability of resting-state activity using two imaging contrasts and two states with two preprocessing strategies. CBF-mean demonstrated the highest reliability, followed by ALFF of BOLD oscillations, which also showed high reliability. CBF-ALFF, BOLD- and CBF-FC showed reasonable reliability, while much lower than those of CBF-mean and BOLD-ALFF. Further, for both imaging contrasts, slightly higher reliability in the EO than the EC state was observed, especially for long-term ICC of BOLD-ALFF in the visual system and both short- and long-term ICC of BOLD-FC in the default mode network.

Test–retest reliability of BOLD and CBF

It is of crucial importance to evaluate the test–retest reliability of brain activity measures and build human brain function studies on these reliable measures (Hodkinson et al., 2013; Jahng et al., 2005; Jann et al., 2014; Kublbock et al., 2014; Marcus, 2014; Patriat et al., 2013; Shehzad et al., 2009; Zhu et al., 2014; Zuo et al., 2010a,b,c, 2013; Zuo and Xing, 2014). In alignment to several previous test–retest reliability studies of average CBF (Hodkinson et al., 2013; Jahng et al.,

Fig. 5. Comparisons of ICC maps between different metrics. Comparisons of ICC maps between CBF-mean and BOLD-ALFF, between CBF-mean and CBF-ALFF, between CBF-mean and BOLD-FC, between CBF-mean and CBF-FC, between BOLD- and CBF-ALFF, between BOLD-ALFF and BOLD-FC, between CBF-ALFF and CBF-FC, and between BOLD- and CBF-FC for short- and long-term ICC in EO (left column) and EC (right column) states were shown. *Z*-axial coordinates in the Talairach and Tournoux space are from −9 to 51 mm in steps of 12 mm.


CBF-mean
BOLD-ALFF
CBF-ALFF
BOLD-FC
CBF-FC
Histogram of short-term ICC
Histogram of long-term ICC
EO
EC
Number of voxels
ICC

2005; Jann et al., 2014), we showed excellent reliability of CBF-mean almost across the whole brain in both EO and EC states. Consistent with previous reliability investigations of BOLD-ALFF in EO (Kublbock et al., 2014; Zuo et al., 2010a), we observed high reliability in the gray matter and most white matter regions and extended the findings to both EO and EC resting states. To date, no study has reported the reliability of voxel-wise CBF-ALFF. Here using 3D pCASL with improved sensitivity, we showed that the reliability of CBF-ALFF to be reasonably high in both EO and EC states, again, higher in the gray matter and lower in the white matter. Consistent with previous reliability studies on BOLD-FC (Patriat et al., 2013; Shehzad et al., 2009), we showed reasonably high reliability within the default mode network which positively correlated with the PCC and part of the task positive network which negatively correlated with the PCC. Previous reliability study on functional connectivity of CBF was based on ICA, which showed high short-term reliability in the PCC and inferior parietal regions of the default mode network in the EO state (Jann et al., 2014). Here we replicated their findings with seed-based CBF-FC and extended that CBF-FC was reliable within the default mode network for both short- and long-term and in both the EO and EC states, though with a trend of decreased reliability in EC.

For the question of which contrast is more reliable, we observed consistently higher short- and long-term ICC of CBF-mean than BOLD-, and CBF-ALFF and BOLD- and CBF-FC. Further, ICC for BOLD-ALFF was significantly higher than CBF-ALFF and BOLD-FC; and ICC of CBF-FC was lower than that of CBF-ALFF and BOLD-FC. Lower reliability of CBF-FC than BOLD-FC using seed-based approach in the current study agreed with the conclusions drawn by Jann and colleagues using ICA (Jann et al., 2015). To note, functional connectivity is a metric reflecting relationships among multi voxels (Zuo and Xing, 2014). It is more straightforward to compare ICC values between BOLD and CBF contrasts both at single voxel level, e.g., among CBF-mean, BOLD- and CBF-ALFF.

Average CBF resulted in the highest reliability, which has important implications for longitudinal studies such as pharmacological modulations. Resting-state CBF with higher reliability would be able to better track the underlying longitudinal changes in brain physiology while reducing potential contaminations by confounding factors. Though lower than CBF-mean, test–retest reliability of BOLD-ALFF was also high, and this warrants its suitability for resting-state fMRI. In general, metrics with moderate to almost perfect test–retest reliability (ICC ≥ 0.4) are considered as reliable for neuroimaging studies (Zuo and Xing, 2014). Here, we showed that ICC values of BOLD-ALFF and CBF-mean were higher than 0.4 almost across all the gray matter voxels. Thus, researchers should be confident to compute BOLD-ALFF and CBF-mean across the gray matter areas. In comparison to BOLD which is a composite reflection of CBF, cerebral brain volume and cerebral metabolic rate of oxygen, CBF measured by ASL could provide a single physiological parameter related to cerebral metabolism and neuronal activity. The complementary characteristics of CBF (high reliability and quantitative) and BOLD (high temporal resolution and signal to noise ratio, reasonable reliability) suggests the combination of these two resting-state fMRI contrasts may offer a powerful tool for characterizing the brain's resting state in future studies. Although across the whole brain, ICC values of BOLD-FC were lower than those of BOLD-ALFF and CBF-mean (Figs. 4–5), within the regions that showed significant correlations with the seed, the ICC values were reasonably high, motivating wide applications of FC. It is not surprising to observe the relatively low reliability for ALFF and FC of CBF, which could be possibly due to low temporal resolution and low signal to noise ratio thus low sensitivity. For those regions with ICC of CBF-ALFF and CBF-FC lower than 0.4, researchers should consider using corresponding metrics based on BOLD contrast. Future studies targeting improving reliability of CBF dynamics should aim at increasing its sensitivity and temporal resolution such as optimized background suppression (Maleki et al., 2012) and labeling schemes (Vidorreta et al., 2013).

ALFF is sensitive to physiological noises in regions near blood vessels and ventricles (Zang et al., 2007; Zou et al., 2008; Zuo et al., 2010a), which is further supported by our findings that the short-term reliability of ALFF was higher than average CBF in these regions. Although advanced preprocessing strategy including ICA noise components removal (Beall and Lowe, 2007, 2010) and additional regression of head motion, CSF and white matter time series (Garrett et al., 2010), could help to eliminate noise contaminations, the observations of residual noises contamination suggested future work on denoising such as physiological noise modeling using simultaneously recorded physiological pulsations and more robust indices for measurement of fluctuation amplitude. Fractional ALFF (fALFF) was proposed to suppress these physiological noises by normalizing ALFF using a ratio between amplitude of low-frequency fluctuations and the amplitude of the overall frequency range. Note that, fALFF was not adopted in the present study due to the following considerations: firstly, as mentioned above, fALFF is mathematically calculated as a ratio, its reliability was demonstrated to be relatively low (Kublbock et al., 2014; Zuo et al., 2010a); secondly, the temporal resolution of CBF time series is low and could only approximately cover the low frequency range of resting-state fluctuations. Thus it is not feasible to calculate fALFF based on CBF time series. To keep consistency between BOLD and CBF contrasts, ALFF instead of fALFF was used as a regional activity metric.

In the current study, we showed that the short-term reliability were consistently higher than those of long-term (Table S2, Fig. S5). It should be noted that time of day might be a potential impacting factor on test–retest reliability, especially for short-term reliability which were calculated based on data from morning and afternoon sessions. Hasher and colleagues have showed time of day effect on within-subject differences of behavioral performance and functional activation in both young and older adults (Anderson et al., 2014; Hogan et al., 2009; May et al., 2005). In addition, both CBF-mean and BOLD-FC had been demonstrated to be decreased moving from morning to afternoon (Hodkinson et al., 2013).

Test–retest reliability of eyes-open and eyes-closed states

Previous test–retest studies on BOLD-ALFF (Kublbock et al., 2014; Zuo et al., 2010a) have been conducted in the EO state. Several test–retest studies have been conducted on CBF-mean, mainly focused on specific ASL imaging techniques using ROI analyses, without specifying the EO or EC state (Chen et al., 2011; Gevers et al., 2011; Jain et al., 2012; Petersen et al., 2010; Wu et al., 2014). Two test–retest studies on CBF-mean were conducted with participants' eyes open (Hodkinson et al., 2013; Wang, 2012) and other two with eyes closed (Huang et al., 2013; Wang et al., 2011a), separately. Only one study acquired CBF data both with participants' eyes open and closed. However, the assessment of test–retest reliability of CBF-mean was calculated based on the average CBF-mean of two resting states (Hermes et al., 2007).

As far as we know, only one resting-state study investigated the test–retest reliability between EO and EC states (Patriat et al., 2013), though focused on resting-state functional connectivity. Patriat and colleagues showed that having subjects keep their eyes open resulted in higher reliability of functional connectivity within the visual, motor, and default mode networks and higher reliability of functional connectivity when examining all the within-network connections including the default mode, attention, auditory, visual and motor networks. Consistent with Patriat and colleagues, we showed that the reliability of default mode network was higher in the EO than in the EC state using BOLD, and extended the finding to CBF based FC. However, it is

Fig. 6. Histograms of short- and long-term ICC between two resting states. Histograms of short- (left column) and long-term (right column) ICC between two resting states using CBF-mean, BOLD- and CBF-ALFF, and BOLD- and CBF-FC.

Table 2
Effect size of comparisons of ICC between eyes open and eyes closed states.

	Short-term	Long-term
CBF-mean	0.06	0.07
BOLD-ALFF	0.10	0.19
CBF-ALFF	0.05	0.06
BOLD-FC	0.01	0.23
CBF-FC	0.16	0.25

Effect size was calculated from Wilcoxon's signed rank test (see Eq. (5)).

still unclear which state would yield more reliable regional activity at single voxel level, including CBF-mean and ALFF of CBF and BOLD. The current results confirmed the previous findings that EO was a slightly more reliable state from the whole brain perspective. This finding also applies to temporal oscillations of multi imaging contrasts besides functional connectivity of BOLD. The convergent results from multiple imaging contrasts and multiple computational indices suggested that future resting-state studies, especially for those focusing on visual system and default mode network, may instruct the participants to keep eyes open, though the effect size between EO and EC was small.

During the EO state, varied amount of visual information input and remarkable reduction of alpha rhythm might result in higher between-subject variability and/or smaller within-subject variability, which may contribute to the observed higher reliability. This was supported by our finding that the primary visual cortex showed the most remarkable state effect on long-term reliability of BOLD-ALFF. In addition, we observed higher long-term reliability of CBF-ALFF in the lateral prefrontal cortex in the EO state, which might be induced by larger inter-subject difference in neuronal activity or by artifacts resulting from eye movement such as blinks. An shown in a previous study, the amplitude of high-frequency (>0.1 Hz) fluctuation in the lateral prefrontal cortex was higher in the EO state (Yuan et al., 2014). The temporal resolution of CBF time series was restricted to the low-frequency range (<0.1 Hz). Therefore those high frequency components would alias into the low-frequency CBF fluctuations, while having less impact low-frequency of BOLD fluctuations.

It is interesting to note that for specific brain regions, there was a trend that the impact of resting states on its reliability seemed to be dependent on the experimental design. For example, there was a trend that short-term ICC of BOLD-ALFF in EO was lower in the motor and auditory cortices than in EC; these regions were frequently reported to feature lower BOLD-ALFF in EO than EC, where EO and EC data were acquired within an hour (Liu et al., 2013; Yuan et al., 2014). However, for the long-term observation, reliability of BOLD-ALFF in the motor areas was changed to be higher in EO, consistent with the findings in the visual system. Moreover, for future studies targeting those regions that did not show significant ICC differences between EO and EC, it would not matter whether to instruct the participants with eyes open or eyes closed.

Beyond EC and EO without fixation, during the resting-state fMRI studies, the participants are sometimes instructed to keep eyes open with fixation. Taglizucchi and colleagues showed that instructing subjects to keep their eyes open with/without fixation during the experiment resulted in a decreased amount of sleep than eyes closed (Tagliazucchi and Laufs, 2014). Furthermore, subjects with eyes open with the fixation showed less sleep than group with eyes open without fixation and eyes closed group, suggesting that eyes open, especially with fixation supports the maintenance of wakefulness in a resting-state setting. In EEG studies, alpha rhythm blockade in the eyes open without fixation is a very solid phenomenon, while the difference between eyes open with fixation and without fixation was seldom reported. In addition, eyes open without fixation and eyes closed are two natural states of human beings. Eyes open with fixation is a more cognitive-demanding state as compared to the eyes closed state and eyes open without fixation state, especially for patients. It would be relatively difficult to ask patients to maintain their gaze on a fixation for quite a few minutes. As did in some previous studies (Liu et al., 2013; Yang et al., 2007; Yuan et al., 2014; Zou et al., 2009; Zou et al., 2015), the current study also only focused on the comparisons of the eyes closed state and eyes open without fixation state and hopefully this study would be more helpful to clinical studies. Future comprehensive test–retest reliability studies including EC and EO with/without fixation are warranted.

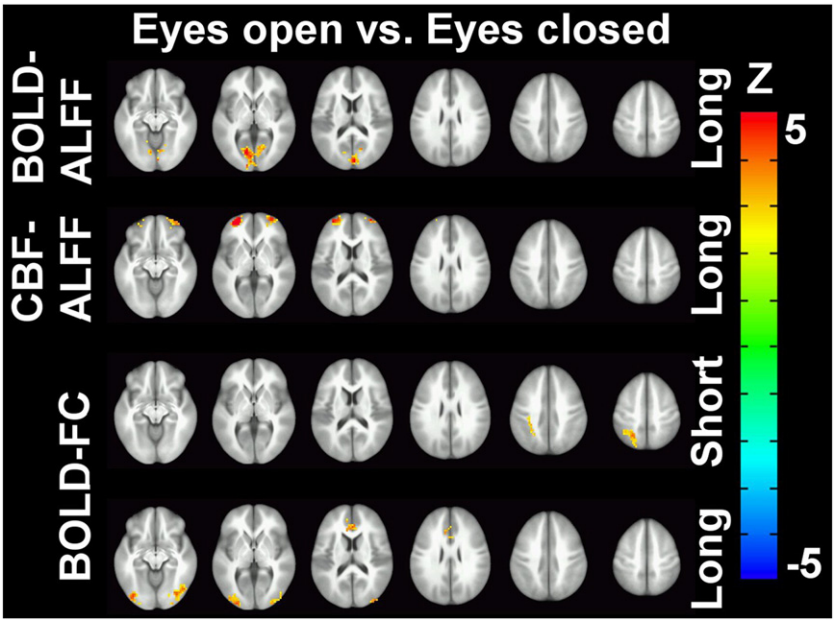


Fig. 7. Comparisons of ICC between two resting states. Significant long-term ICC differences between eyes open and eyes closed states were observed using BOLD- and CBF-ALFF. Significant ICC differences between two resting states were also observed using BOLD-FC for both short- and long-term. *Z*-axial coordinates in the Talairach and Tournoux space are from −9 to 51 mm in steps of 12 mm.

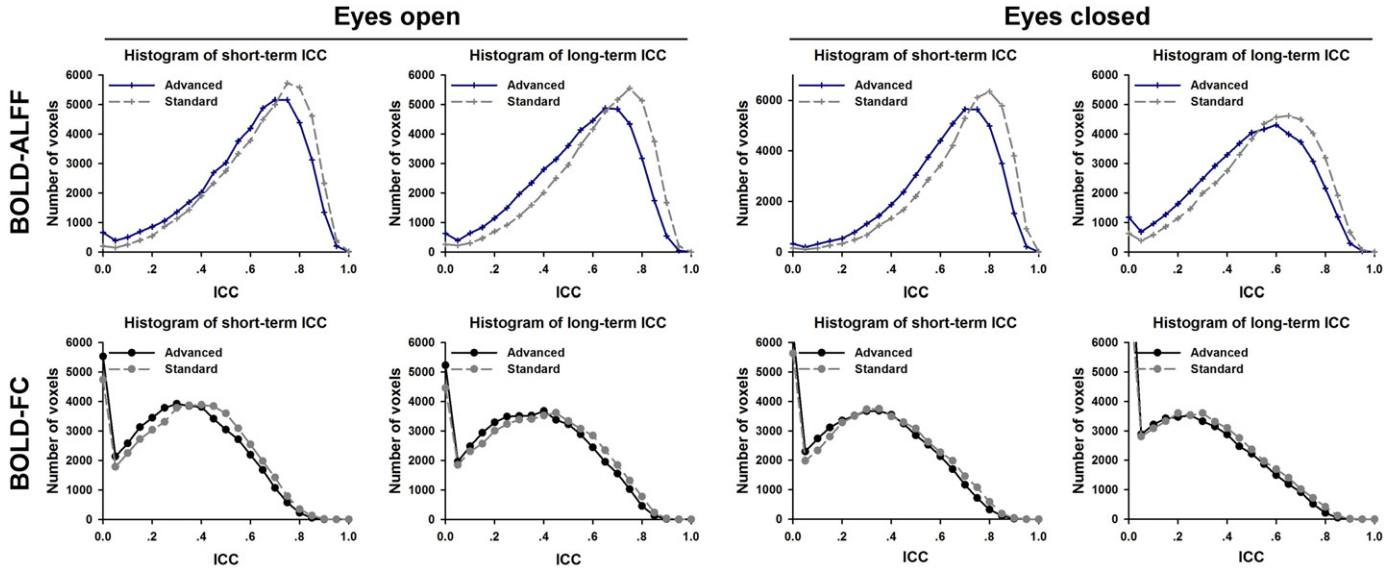


Fig. 8. Histograms of ICC between advanced and standard preprocessing in two resting states. Histograms of short- and long-term ICC between advanced and standard preprocessing of BOLD-ALFF, and BOLD-FC metrics in two resting states.

Effect of preprocessing on test–retest reliability

To minimize the potential contaminations of physiological pulsations and head motion, we employed extended data preprocessing strategies including ICA denoising (Beall and Lowe, 2007, 2010), regression of covariates of no interests including head motion parameters, mean signals within the white matter and CSF masks (Garrett et al., 2010). After denoising, test–retest reliability of BOLD-ALFF decreased (Table 3 and Fig. 8), which was consistent with prior work based on BOLD emphasizing the susceptibility of ALFF to physiological noises (Zang et al., 2007; Zuo et al., 2010a) as well as previous findings that motion regression decreased reliability of ALFF (Yan et al., 2013a). Similar to our findings of ALFF, test–retest reliability of BOLD-FC was also decreased with denoising (Table 3 and Fig. 8), consistent with the observation by Birn and colleagues (Birn et al., 2014) that reliability of BOLD-FC was decreased after physiological noise correction.

The contributing effect of noises on test–retest reliability of both ALFF and FC was not surprising. It was shown that physiological pulsations (Birn et al., 2014) and motion showed moderate reliability (Van Dijk et al., 2012; Zuo et al., 2014), which might inflate the reliability of ALFF and FC. Although physiological noise and motion corrections reduced the reliability and variability of ALFF and FC between subjects, it might increase the power of group-level analyses, such as boosting the predictability of age using ALFF (Garrett et al., 2010) and leading to more accurate estimates of group maps of FC (Birn et al., 2014).

Conclusions and future directions

To conclude, both BOLD oscillations and average CBF showed high test–retest reliability across the whole brain, and BOLD-FC showed reasonably high reliability within the regions showing significant correlation with the seed, which would be preferable for clinical applications and longitudinal observations. Further, resting-state fMRI in the EO state was slightly more reliable than in the EC state though the effect size was small. We recommend to collect both BOLD and pCASL resting-state data in the EO state in future studies. The observation of lower test–retest reliability of CBF-ALFF and CBF-FC calls for further improvements in signal-to-noise ratio of ASL imaging techniques (Li et al., 2015).

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.07.044.

Table 3
Effect size of the difference of ICC between standard and advanced preprocessing strategies.

	Eyes open		Eyes closed	
	Short-term	Long-term	Short-term	Long-term
BOLD-ALFF	0.22	0.32	0.32	0.24
BOLD-FC	0.18	0.15	0.14	0.16

Effect size was calculated from Wilcoxon's signed rank test (see Eq. (5)).

Acknowledgments

This work was supported by the China's National Strategic Basic Research Program (973) (2012CB720700 and 2015CB856400), the Natural Science Foundation of China (81201142, 81227003, 30921064, 91132302, 90820307), and the Ministry of Science and Technology of China grant (2012CB825500).

References

Anderson, J.A., Campbell, K.L., Amer, T., Grady, C.L., & Hasher, L., 2014. Timing is everything: age differences in the cognitive control network are modulated by time of day. Psychol. Aging 29, 648–657.
Beall, E.B., & Lowe, M.J., 2007. Isolating physiologic noise sources with independently determined spatial measures. Neuroimage 37, 1286–1300.
Beall, E.B., & Lowe, M.J., 2010. The non-separability of physiologic noise in functional connectivity MRI with spatial ICA at 3 T. J. Neurosci. Methods 191, 263–276.
Berger, H., 1929. U¨ ber das Elektrenkephalogramm des Menschen. Arch. Psychiatr. Nervenkr. 87, 44.
Berger, H., 1930. U¨ ber das Elektrenkephalogramm des Menschen II. J. Psychol. Neurol. 40, 20.
Birn, R.M., Cornejo, M.D., Molloy, E.K., Patriat, R., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., & Prabhakaran, V., 2014. The influence of physiological noise correction on test–retest reliability of resting-state functional connectivity. Brain Connect. 4, 511–522.
Biswal, B.B., Mennes, M., Zuo, X.N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S., Buckner, R.L., Colcombe, S., Dogonowski, A.M., Ernst, M., Fair, D., Hampson, M., Hoptman, M.J., Hyde, J.S., Kiviniemi, V.J., Kotter, R., Li, S.J., Lin, C.P., Lowe, M.J., Mackay, C., Madden, D.J., Madsen, K.H., Margulies, D.S., Mayberg, H.S., McMahon, K., Monk, C.S., Mostofsky, S.H., Nagel, B.J., Pekar, J.J., Peltier, S.J., Petersen, S.E., Riedl, V., Rombouts, S.A., Rypma, B., Schlaggar, B.L., Schmidt, S., Seidler, R.D., Siegle, G.J., Sorg, C., Teng, G.J., Veijola, J., Villringer, A., Walter, M., Wang, L., Weng, X.C., Whitfield-Gabrieli, S., Williamson, P., Windischberger, C., Zang, Y.F., Zhang, H.Y., Castellanos, F.X., & Milham, M.P., 2010. Toward discovery science of human brain function. Proc. Natl. Acad. Sci. U. S. A. 107, 4734–4739.
Chen, Y., Wang, D.J., & Detre, J.A., 2011. Test–retest reliability of arterial spin labeling with common labeling strategies. J. Magn. Reson. Imaging 33, 940–949.

Chuang, K.H., van Gelderen, P., Merkle, H., Bodurka, J., Ikonomidou, V.N., Koretsky, A.P., Duyn, J.H., & Talagala, S.L., 2008. Mapping resting-state functional connectivity using perfusion MRI. Neuroimage 40, 1595–1605.
Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.
Dai, W., Garcia, D., de Bazelaire, C., & Alsop, D.C., 2008. Continuous flow-driven inversion for arterial spin labeling using pulsed radio frequency and gradient fields. Magn. Reson. Med. 60, 1488–1497.
Garrett, D.D., Kovacevic, N., McIntosh, A.R., & Grady, C.L., 2010. Blood oxygen level-dependent signal variability is more than just noise. J. Neurosci. 30, 4914–4921.
Garrett, D.D., Kovacevic, N., McIntosh, A.R., & Grady, C.L., 2011. The importance of being variable. J. Neurosci. 31, 4496–4503.
Garrett, D.D., Samanez-Larkin, G.R., MacDonald, S.W., Lindenberger, U., McIntosh, A.R., & Grady, C.L., 2013. Moment-to-moment brain signal variability: a next frontier in human brain mapping? Neurosci. Biobehav. Rev. 37, 610–624.
Gevers, S., van Osch, M.J., Bokkers, R.P., Kies, D.A., Teeuwisse, W.M., Majoie, C.B., Hendrikse, J., & Nederveen, A.J., 2011. Intra- and multicenter reproducibility of pulsed, continuous and pseudo-continuous arterial spin labeling methods for measuring cerebral perfusion. J. Cereb. Blood Flow Metab. 31, 1706–1715.
Han, Y., Wang, J., Zhao, Z., Min, B., Lu, J., Li, K., He, Y., & Jia, J., 2011. Frequency-dependent changes in the amplitude of low-frequency fluctuations in amnestic mild cognitive impairment: a resting-state fMRI study. Neuroimage 55, 287–295.
Hermes, M., Hagemann, D., Britz, P., Lieser, S., Rock, J., Naumann, E., & Walter, C., 2007. Reproducibility of continuous arterial spin labeling perfusion MRI after 7 weeks. MAGMA 20, 103–115.
Hodkinson, D.J., Krause, K., Khawaja, N., Renton, T.F., Huggins, J.P., Vennart, W., Thacker, M.A., Mehta, M.A., Zelaya, F.O., Williams, S.C., & Howard, M.A., 2013. Quantifying the test–retest reliability of cerebral blood flow measurements in a clinical model of on-going post-surgical pain: a study using pseudo-continuous arterial spin labelling. Neurol. Clin. 3, 301–310.
Hogan, M.J., Kelly, C.A., Verrier, D., Newell, J., Hasher, L., & Robertson, I.H., 2009. Optimal time-of-day and consolidation of learning in younger and older adults. Exp. Aging Res. 35, 107–128.
Huang, D., Wu, B., Shi, K., Ma, L., Cai, Y., & Lou, X., 2013. Reliability of three-dimensional pseudo-continuous arterial spin labeling MR imaging for measuring visual cortex perfusion on two 3 T scanners. PLoS One 8, e79471.
Jahng, G.H., Song, E., Zhu, X.P., Matson, G.B., Weiner, M.W., & Schuff, N., 2005. Human brain: reliability and reproducibility of pulsed arterial spin-labeling perfusion MR imaging. Radiology 234, 909–916.
Jain, V., Duda, J., Avants, B., Giannetta, M., Xie, S.X., Roberts, T., Detre, J.A., Hurt, H., Wehrli, F.W., & Wang, D.J., 2012. Longitudinal reproducibility and accuracy of pseudo-continuous arterial spin-labeled perfusion MR imaging in typically developing children. Radiology 263, 527–536.
Jann, K., Gee, D.G., Kilroy, E., Schwab, S., Smith, R.X., Cannon, T.D., & Wang, D.J., 2014. Functional connectivity in BOLD and CBF data: similarity and reliability of resting brain networks. Neuroimage 106C, 111–122.
Jann, K., Gee, D.G., Kilroy, E., Schwab, S., Smith, R.X., Cannon, T.D., & Wang, D.J., 2015. Functional connectivity in BOLD and CBF data: similarity and reliability of resting brain networks. Neuroimage 106, 111–122.
Jao, T., Vertes, P.E., Alexander-Bloch, A.F., Tang, I.N., Yu, Y.C., Chen, J.H., & Bullmore, E.T., 2013. Volitional eyes opening perturbs brain dynamics and functional connectivity regardless of light input. Neuroimage 69, 21–34.
Kublbock, M., Woletz, M., Hoflich, A., Sladky, R., Kranz, G.S., Hoffmann, A., Lanzenberger, R., & Windischberger, C., 2014. Stability of low-frequency fluctuation amplitudes in prolonged resting-state fMRI. Neuroimage 103C, 249–257.
Li, X., Wang, D., Auerbach, E.J., Moeller, S., Ugurbil, K., & Metzger, G.J., 2015. Theoretical and experimental evaluation of multi-band EPI for high-resolution whole brain pCASL imaging. Neuroimage 106, 170–181.
Liang, X., Zou, Q., He, Y., & Yang, Y., 2013. Coupling of functional connectivity and regional cerebral blood flow reveals a physiological basis for network hubs of the human brain. Proc. Natl. Acad. Sci. U. S. A. 110, 1929–1934.
Liu, D., Dong, Z., Zuo, X., Wang, J., & Zang, Y., 2013. Eyes-open/eyes-closed dataset sharing for reproducibility evaluation of resting state FMRI data analysis methods. Neuroinformatics 11, 469–476.
Lui, S., Huang, X., Chen, L., Tang, H., Zhang, T., Li, X., Li, D., Kuang, W., Chan, R.C., Mechelli, A., Sweeney, J.A., & Gong, Q., 2009. High-field MRI reveals an acute impact on brain function in survivors of the magnitude 8.0 earthquake in China. Proc. Natl. Acad. Sci. U. S. A. 106, 15412–15417.
Maleki, N., Dai, W., & Alsop, D.C., 2012. Optimization of background suppression for arterial spin labeling perfusion imaging. MAGMA 25, 127–133.
Marcus, E., 2014. Credibility and reproducibility. Cancer Cell 26, 771–772.
May, C.P., Hasher, L., & Foong, N., 2005. Implicit memory, age, and time of day: paradoxical priming effects. Psychol. Sci. 16, 96–100.
McGraw, K.O., & Wong, S., 1996. Forming inferences about some intraclass correlation coefficients. Psychol. Methods 1, 30.
Patriat, R., Molloy, E.K., Meier, T.B., Kirk, G.R., Nair, V.A., Meyerand, M.E., Prabhakaran, V., & Birn, R.M., 2013. The effect of resting condition on resting-state fMRI reliability and consistency: a comparison between resting with eyes open, closed, and fixated. Neuroimage 78, 463–473.
Petersen, E.T., Mouridsen, K., Golay, X., & all named co-authors of the, Q.t.-r.s., 2010. The QUASAR reproducibility study, part II: results from a multi-center arterial spin labeling test–retest study. Neuroimage 49, 104–113.
Searle, S.R., Casella, G., & McCulloch, C.E., 1992. Variance Components. John Wiley & Sons, INC., New York.
Shehzad, Z., Kelly, A.M., Reiss, P.T., Gee, D.G., Gotimer, K., Uddin, L.Q., Lee, S.H., Margulies, D.S., Roy, A.K., Biswal, B.B., Petkova, E., Castellanos, F.X., & Milham, M.P., 2009. The resting brain: unconstrained yet reliable. Cereb. Cortex 19, 2209–2229.
Shrout, P.E., & Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. Psychol. Bull. 86, 420–428.
Smith, S.M., Vidaurre, D., Beckmann, C.F., Glasser, M.F., Jenkinson, M., Miller, K.L., Nichols, T.E., Robinson, E.C., Salimi-Khorshidi, G., Woolrich, M.W., Barch, D.M., Ugurbil, K., & Van Essen, D.C., 2013. Functional connectomics from resting-state fMRI. Trends Cogn. Sci. 17, 666–682.
Tagliazucchi, E., & Laufs, H., 2014. Decoding wakefulness levels from typical fMRI resting-state data reveals reliable drifts between wakefulness and sleep. Neuron 82, 695–708.
Van Dijk, K.R., Sabuncu, M.R., & Buckner, R.L., 2012. The influence of head motion on intrinsic functional connectivity MRI. Neuroimage 59, 431–438.
Vidorreta, M., Wang, Z., Rodriguez, I., Pastor, M.A., Detre, J.A., & Fernandez-Seara, M.A., 2013. Comparison of 2D and 3D single-shot ASL perfusion fMRI sequences. Neuroimage 66, 662–671.
Wang, Z., 2012. Improving cerebral blood flow quantification for arterial spin labeled perfusion MRI by removing residual motion artifacts and global signal fluctuations. Magn. Reson. Imaging 30, 1409–1415.
Wang, Y., Saykin, A.J., Pfeuffer, J., Lin, C., Mosier, K.M., Shen, L., Kim, S., & Hutchins, G.D., 2011a. Regional reproducibility of pulsed arterial spin labeling perfusion imaging at 3 T. Neuroimage 54, 1188–1195.
Wang, Z., Yan, C., Zhao, C., Qi, Z., Zhou, W., Lu, J., He, Y., & Li, K., 2011b. Spatial patterns of intrinsic brain activity in mild cognitive impairment and Alzheimer's disease: a resting-state functional MRI study. Hum. Brain Mapp. 32, 1720–1740.
Wu, W.C., Fernandez-Seara, M., Detre, J.A., Wehrli, F.W., & Wang, J., 2007. A theoretical and experimental investigation of the tagging efficiency of pseudocontinuous arterial spin labeling. Magn. Reson. Med. 58, 1020–1027.
Wu, B., Lou, X., Wu, X., & Ma, L., 2014. Intra- and interscanner reliability and reproducibility of 3D whole-brain pseudo-continuous arterial spin-labeling MR perfusion at 3 T. J. Magn. Reson. Imaging 39, 402–409.
Yan, C., Liu, D., He, Y., Zou, Q., Zhu, C., Zuo, X., Long, X., & Zang, Y., 2009. Spontaneous brain activity in the default mode network is sensitive to different resting-state conditions with limited cognitive load. PLoS One 4, e5743.
Yan, C.G., Cheung, B., Kelly, C., Colcombe, S., Craddock, R.C., Di Martino, A., Li, Q., Zuo, X.N., Castellanos, F.X., & Milham, M.P., 2013a. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. Neuroimage 76, 183–201.
Yan, C.G., Craddock, R.C., Zuo, X.N., Zang, Y.F., & Milham, M.P., 2013b. Standardizing the intrinsic brain: towards robust measurement of inter-individual variation in 1000 functional connectomes. Neuroimage 80, 246–262.
Yang, H., Long, X.Y., Yang, Y., Yan, H., Zhu, C.Z., Zhou, X.P., Zang, Y.F., & Gong, Q.Y., 2007. Amplitude of low frequency fluctuation within visual areas revealed by resting-state functional MRI. Neuroimage 36, 144–152.
Yu, R., Chien, Y.L., Wang, H.L., Liu, C.M., Liu, C.C., Hwang, T.J., Hsieh, M.H., Hwu, H.G., & Tseng, W.Y., 2014. Frequency-specific alternations in the amplitude of low-frequency fluctuations in schizophrenia. Hum. Brain Mapp. 35, 627–637.
Yuan, B.K., Wang, J., Zang, Y.F., & Liu, D.Q., 2014. Amplitude differences in high-frequency fMRI signals between eyes open and eyes closed resting states. Front. Hum. Neurosci. 8, 503.
Zang, Y.F., He, Y., Zhu, C.Z., Cao, Q.J., Sui, M.Q., Liang, M., Tian, L.X., Jiang, T.Z., & Wang, Y.F., 2007. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. Brain Dev. 29, 83–91.
Zhu, L., Fan, Y., Zou, Q., Wang, J., Gao, J.H., & Niu, Z., 2014. Temporal reliability and lateralization of the resting-state language network. PLoS One 9, e85880.
Zou, Q.H., Zhu, C.Z., Yang, Y., Zuo, X.N., Long, X.Y., Cao, Q.J., Wang, Y.F., & Zang, Y.F., 2008. An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. J. Neurosci. Methods 172, 137–141.
Zou, Q., Long, X., Zuo, X., Yan, C., Zhu, C., Yang, Y., Liu, D., He, Y., & Zang, Y., 2009. Functional connectivity between the thalamus and visual cortex under eyes closed and eyes open conditions: a resting-state fMRI study. Hum. Brain Mapp. 30, 3066–3078.
Zou, Q., Gu, H., Wang, D.J., Gao, J.H., & Yang, Y., 2011. Quantification of load dependent brain activity in parametric N-back working memory tasks using pseudo-continuous arterial spin labeling (pCASL) perfusion imaging. J. Cogn. Sci.(Seoul) 12, 127–210.
Zou, Q., Ross, T.J., Gu, H., Geng, X., Zuo, X.N., Hong, L.E., Gao, J.H., Stein, E.A., Zang, Y.F., & Yang, Y., 2013. Intrinsic resting-state activity predicts working memory brain activation and behavioral performance. Hum. Brain Mapp. 34, 3204–3215.
Zou, Q., Yuan, B.K., Gu, H., Liu, D., Wang, D.J., Gao, J.H., Yang, Y., & Zang, Y.F., 2015. Detecting static and dynamic differences between eyes-closed and eyes-open resting states using ASL and BOLD fMRI. PLoS One 10, e0121757.
Zuo, X.N., & Xing, X.X., 2014. Test–retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: a systems neuroscience perspective. Neurosci. Biobehav. Rev. 45, 100–118.
Zuo, X.N., Di Martino, A., Kelly, C., Shehzad, Z.E., Gee, D.G., Klein, D.F., Castellanos, F.X., Biswal, B.B., & Milham, M.P., 2010a. The oscillating brain: complex and reliable. Neuroimage 49, 1432–1445.
Zuo, X.N., Kelly, C., Adelstein, J.S., Klein, D.F., Castellanos, F.X., & Milham, M.P., 2010b. Reliable intrinsic connectivity networks: test–retest evaluation using ICA and dual regression approach. Neuroimage 49, 2163–2177.
Zuo, X.N., Kelly, C., Di Martino, A., Mennes, M., Margulies, D.S., Bangaru, S., Grzadzinski, R., Evans, A.C., Zang, Y.F., Castellanos, F.X., & Milham, M.P., 2010c. Growing together and growing apart: regional and sex differences in the lifespan developmental trajectories of functional homotopy. J. Neurosci. 30, 15034–15043.
Zuo, X.N., Xu, T., Jiang, L., Yang, Z., Cao, X.Y., He, Y., Zang, Y.F., Castellanos, F.X., & Milham, M.P., 2013. Toward reliable characterization of functional homogeneity in the human brain: preprocessing, scan duration, imaging resolution and computational space. Neuroimage 65, 374–386.

Zuo, X.-N., Anderson, J.S., Bellec, P., Birn, R.M., Biswal, B.B., Blautzik, J., Breitner, J.C.S., Buckner, R.L., Calhoun, V.D., Castellanos, F.X., Chen, A., Chen, B., Chen, J., Chen, X., Colcombe, S.J., Courtney, W., Craddock, R.C., Di Martino, A., Dong, H.-M., Fu, X., Gong, Q., Gorgolewski, K.J., Han, Y., He, Y., He, Y., Ho, E., Holmes, A., Hou, X.-H., Huckins, J., Jiang, T., Jiang, Y., Kelley, W., Kelly, C., King, M., LaConte, S.M., Lainhart, J.E., Lei, X., Li, H.-J., Li, K., Li, K., Lin, Q., Liu, D., Liu, J., Liu, X., Liu, Y., Lu, G., Lu, J., Luna, B., Luo, J., Lurie, D., Mao, Y., Margulies, D.S., Mayer, A.R., Meindl, T., Meyerand, M.E., Nan, W., Nielsen, J.A., O'Connor, D., Paulsen, D., Prabhakaran, V., Qi, Z., Qiu, J., Shao, C., Shehzad, Z., Tang, W., Villringer, A., Wang, H., Wang, K., Wei, D., Wei, G.-X., Weng, X.-C., Wu, X., Xu, T., Yang, N., Yang, Z., Zang, Y.-F., Zhang, L., Zhang, Q., Zhang, Z., Zhang, Z., Zhao, K., Zhen, Z., Zhou, Y., Zhu, X.-T., & Milham, M.P., 2014. An open science resource for establishing reliability and reproducibility in functional connectomics. Sci. Data 1.